AD_____

Award Number: W81XWH-06-1-0331

TITLE: A Novel Strategy for Isolation, Molecular and Functional Characterization of Embryonic Mammary Stem Cells Using Molecular Genetics and Microfluidic Sorting

PRINCIPAL INVESTIGATOR: Geoffrey Wahl, Ph.D.

CONTRACTING ORGANIZATION: The Salk Institute for Biological Studies
La Jolla, CA 92037-1099

REPORT DATE: June 2007

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
01-06-2007	Annual	1 Jun 2006 – 31 May 2007

4. TITLE AND SUBTITLE

A Novel Strategy for Isolation, Molecular and Functional Characterization of Embryonic Mammary Stem Cells Using Molecular Genetics and Microfluidic Sorting

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-06-1-0331

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Geoffrey Wahl, Ph.D.

E-Mail: Wahl@salk.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

The Salk Institute for Biological Studies
La Jolla, CA 92037-1099

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES – Original contains colored plates: ALL DTIC reproductions will be in black and white.

14. ABSTRACT

We are developing a genetic system to identify, isolate, and characterize mammary stem cells. Our system consists of both activator and reporter components. The activator component is dependent both on Wnt signaling, which is essential for mammary gland development, and on doxycycline to toggle the system on and off. The reporter component labels cells for direct visualization. We used a modular design to enable the system to be applied to cancer models and other organs. Data obtained over the past year with a related proof of concept system showed that mammary glands were brightly labeled by H2BGFP fluorescence in response to the Tet transactivator. This initial evaluation demonstrates the feasibility of the approach in vivo. In vitro analyses showed that chromatin is labeled by H2BGFP regardless of cell cycle phase. The molecular reagents and strategies we are developing have broad applications for studies examining the relationship between normal and cancer stem cells, and determining whether they share the same origin.

15. SUBJECT TERMS
Mammary stem cells, cancer stem cells, Wnt signaling

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			USAMRMC
U	U	U	UU	9	**19b. TELEPHONE NUMBER** *(include area code)*

Table of Contents

	Page
Introduction	4
Body	4
Key Research Accomplishments	8
Reportable Outcomes	8
Conclusion	9
References	9
Appendices	none
Supporting Data	none

INTRODUCTION

Studies over the past 15 years indicate that not all cells within a tumor are equivalent when it comes to their ability to regenerate a tumor in nude mice. Rather, only a small subset of cells within the cancer is tumorigenic (1). These cells appear to have self-renewal capacity, and can generate the cellular complexity evident in the original tumor. As these are qualities shared by tissue-specific stem cells, these tumorigenic progenitors have been called "cancer stem cells". Since in some cases the "cancer stem cells" have markers also found on tissue-specific stem cells, it has been suggested that cancer stem cells may originate from adult stem cells or mutated progenitor cells that regain self-renewal potential (2). If cancer stem cells are mutated tissue stem cells, then it will also be important to determine whether they have retained other properties reported for some stem cells, such as infrequent entry into the cell cycle and expression of efflux pumps that could render them refractory to commonly used chemotherapies. It is essential to gain a better understanding of cancer stem cells to answer these critical questions. To achieve this goal, it is necessary to have a general method to isolate and characterize tissue-specific stem cells. The overarching goal of this proposal is to develop a system to identify mammary stem cells in mice with a long-term goal of determining whether they contribute to breast cancer. The proposed approach involves fluorescent labeling of the stem cells within the embryonic mammary gland as we hypothesize that the concentration of stem cells will be higher in the embryo than in the adult due to the presence of fewer cell types early in development. The embryonic mammary stem cells will be isolated using a combination of a powerful molecular genetic system that enables *in vivo* imaging of the putative stem cell population through labeling with a fluorescent protein and a specialized cell sorter capable of sorting small populations. Upon purification, the expression pattern of stem cell genes will be determined. *In vitro* and *in vivo* assays will be used to evaluate self-renewal and multi-lineage differentiation potential, which are the hallmarks of stem cells.

BODY

Task 1: Determine whether there is a population of label retaining cells formed during early mammary gland development

During the May 2006 - June 2007 funding period, we generated five transgenic founder mice carrying the TOP-Tet^{off}VP16 (TOP-tTA) transgene. We determined that mammary buds were not fluorescently labeled in embryos derived from crossing these five founders with TRE-H2BGFP reporter mice. Instead, we observed diffuse fluorescence throughout the embryo (Figure 1). We next investigated why the fluorescence in Wnt-responsive tissues was not detected in order to develop a new strategy to overcome these problems. First, we asked whether the tTA activator component we utilized was capable of activating the TRE-H2BGFP reporter transgene. We introduced a retrovirus encoding tTA driven by the CMV promoter into MEFs isolated from TRE-H2BGFP mice. We observed significant H2BGFP expression in the infected fraction of the population using FACS (Figure 2) and QPCR analyses (not shown). This experiment demonstrated that when expressed at sufficient levels, the tTA activator could stimulate adequate TRE-H2BGFP for detection of the fluorescent cells.

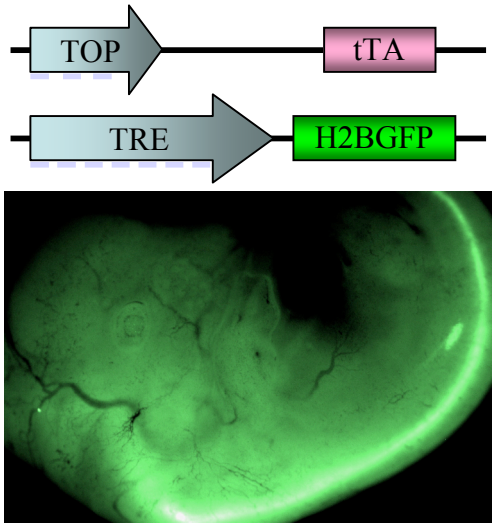


Figure 1: Double transgenic embryos gave nonspecific, diffuse fluorescence

We next determined whether non-dividing cells incorporate H2BGFP into their chromatin, as it may be that in early embryonic mammary development there is little proliferation. We used cell lines in which H2BGFP could be induced by doxycycline administration for this purpose. The cells were arrested in G1 by serum deprivation and mimosine treatment, and doxycycline was then administered for various times, and the intensity of labeling

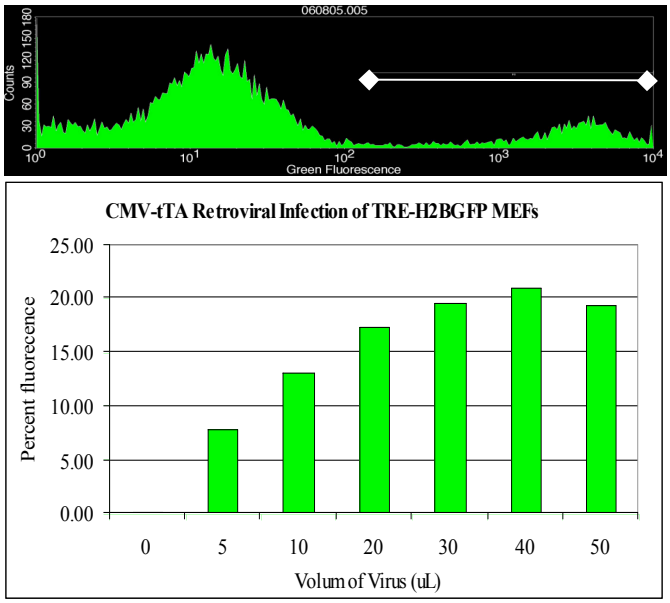


Figure 2: Percentage of cells that are fluorescent (indicated by the gated histogram) after retroviral delivery of tTA to TRE-H2BGFP Mouse Embryonic Fibroblasts

relative to exponentially growing controls determined by FACS and microscopic analyses. Robust and nearly equivalent labeling was observed in growing and arrested cells (Figure 3). The top panel in figure 3 depicts the fluorescence of cycling cells after 48 hours of H2BGFP induction. The middle and lower panels reflect arrest in G_1 and G_0, respectively. The data show that the arrested cells exhibit similar brightness as the cycling cells. The bottom panel shows a second peak representing cells that are not expressing H2BGFP, most likely due to *in vitro* nutrient deprivation. This should not be a concern *in vivo* since all cells should be exposed to adequate nutrients. This experiment demonstrates the feasibility of using H2BGFP labeling to detect cells regardless of cycling.

The most likely explanations for the failure of the original design are the following three factors. First, random insertion of the transgene might compromise expression from the relatively weak TOP promoter if it were sensitive to genomic position effects. Careful analysis of the literature, as well as correspondence with investigators carrying out similar lines of research, revealed that fewer than one in 20 insertion sites enables expression of the TOP promoter we employed. As we analyzed only five founders, we had a low probability of finding one that would express in a Wnt-responsive fashion. The second limitation is the strength of the TOP promoter itself. This promoter consists of three repeats of the deduced minimal Wnt-responsive DNA elements. The third potential limitation comes from the tTA tet transactivator we employed. The original version of tTA that we used is now known to contain a fragment of the VP16 transactivator that is toxic to cells when expressed at high levels. This could clearly select for expression of the transgene at levels that are insufficient to elicit activation of the tetracycline response element promoter (TRE).

To overcome the above known and predicted problems, we completely redesigned the approach to be used as well as the molecular components of the system (Figure 4). The new approach utilizes site-specific recombination into predetermined genomic loci that are known to be expressed ubiquitously, and are constitutively in an open chromatin conformation. This will prevent position and copy number variegation between founder lines, such that all founders should be identical. This will both limit the need to screen numerous founders to detect one that expresses in the predicted Wnt-responsive fashion, and increase the reproducibility of the analysis over time. We have chosen the ROSA26 and ColA1 loci for this purpose as they have been extensively analyzed and validated in the past. We initially chose the ROSA26 locus as it has been used historically for robust and ubiquitous transgene expression, and is readily targeted by homologous recombination (3). We had also developed an effective strategy that utilizes the Cre site-specific recombinase to swap endogenous sequences for any desired transgenic sequence (4). We began to collaborate with a lab that had targeted the ROSA26 locus previously, and provided them with the targeting vector to enable Cre-specific transgene swaps. In return, they provided us targeted ES cells. However, upon analysis of the cells, we found that they were contaminated with mycoplasma, rendering them useless for the production of transgenic mice (data not shown). A central and important goal

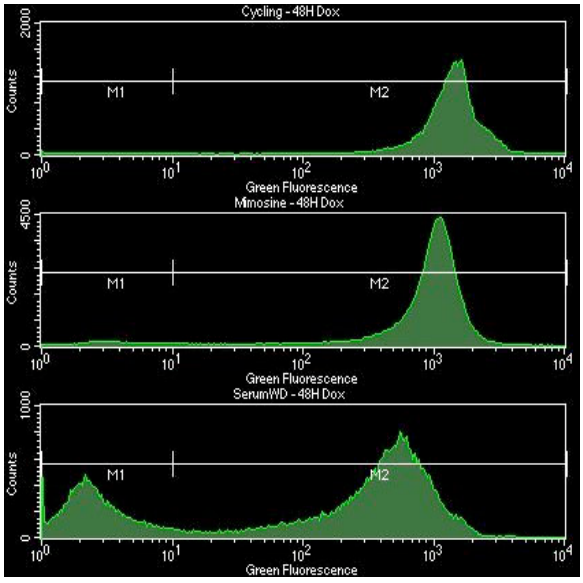


Figure 3: FACS histograms depicting the brightness and degree of H2BGFP expression

of the next year will be to use the targeting vector we made to target the ROSA26 locus ourselves using ES cells we know to be suitable for production of mice. Using funding from another source, we are in parallel pursuing studies at the ColA1 locus (5). This system, developed by the Jaenisch lab at MIT, allows double positive selection into the 3'-untranslated region of the Collagen A1 gene. We have obtained the relevant ES cells, shown them to be mycoplasma free, and are engineering the targeting vector for transgene insertion. It is essential to study two loci, as it is unclear which, if either will enable reliable and robust Wnt-dependent gene expression. If both prove useful, then it will be possible to introduce Wnt-responsive reporters into one, while using the other for introduction of Wnt-activated mutant alleles, or regulated microRNAs, to manipulate Wnt or other developmentally relevant pathways to determine which are involved in mammary stem cell specification. This type of dual locus system also should provide us with future opportunities for tracking cells, as well as introducing potential oncogenic stimuli to follow the course of tumor initiation, progression, and response to treatment.

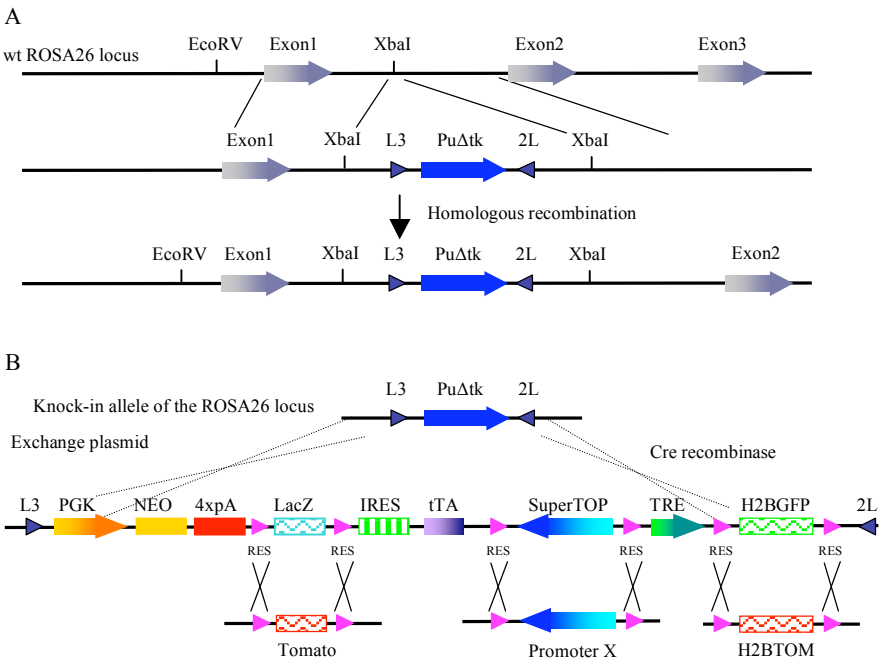
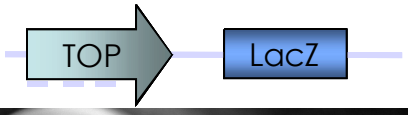


Figure 4: Our targeting strategies. A. For the insertion of the floxed PuΔtk transgene into the ROSA26 locus. B. For the insertion of the SuperTOP-tTA/TRE-H2BGFP transgene into the modified ROSA26 locus.

The following section describes the design of the new transgene cassette we will utilize to optimize Wnt-responsiveness and to minimize transactivator toxicity. We have designed the system to be modular to facilitate replacement of each component such that multiple signaling pathways can be studied in the future, and to take advantage of advancements in technology as new activators or reporters become available. We now use the SuperTOP Wnt-responsive element containing seven TCF binding sites, which has been reported to have 100-fold greater inducibility (Randy Moon, personal communication) than the three TCF sites in the TOP promoter we originally used. We replaced the original tTA with tTA2. tTA2 has the same activation level as the original, but is tolerated at 3-fold higher levels without toxicity (6, 7). We have also included an internal ribosome entry site (IRES) downstream of the transactivator, and have placed either a lacZ gene or a red-fluorescent protein gene, Tomato, adjacent to the IRES. This innovative design will readily enable us to determine whether the transgene is responsive to Wnt signaling, since a response will be revealed by a positive blue stain for LacZ or by the cells fluorescing red. The LacZ was chosen to enable us to determine whether our system is responding with the same sensitivity and tissue specificity as the previously published TOP-Gal mice (8). The cloning of SuperTOP-tTA2-IRES-LacZ has been completed. Due to the problems encountered with the ROSA26 ES cells, we have decided to first target the transgene into the ColA1 locus and test its Wnt responsiveness. To this end, we will target the ColA1 locus, and then test the responsiveness of the ES cells to recombinant Wnt3a by determining whether the ES cells turn blue, and whether they



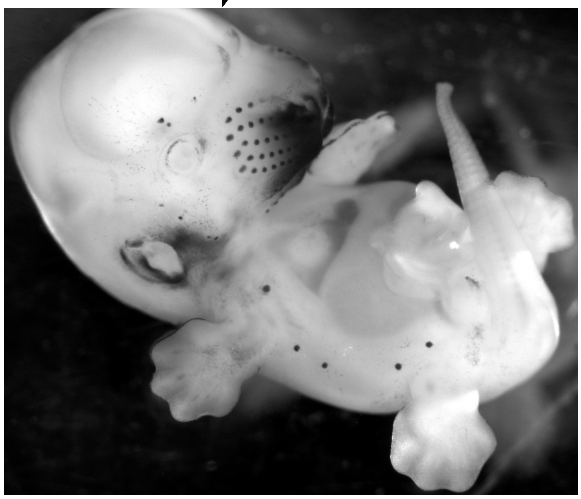


Figure 5: X-Gal staining on E12.5 TOPGAL embryo

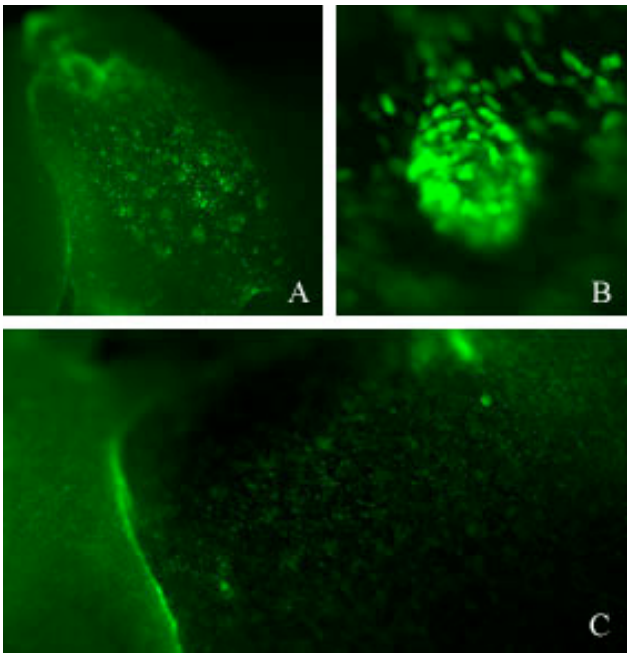


Figure 6A: Vibrissae **6B:** Mammary bud no. 4 **6C:** Mammary buds 3 and 4

express endogenous Wnt responsive genes, such as Axin2 or cyclin D, to levels higher than observed in the absence of inducer. We will also generate mice from the targeted cells using blastocyst injection. We will then determine Wnt3a responsiveness of whole embryos, and compare the staining patterns to those observed for the published TOP-GAL mice. We have found the whole embryo method of evaluating Wnt responsiveness to be robust and sensitive, as shown by the mammary bud-specific staining we readily detect in TOP-Gal mice analyzed at embryonic day 12.5 (Figure 5). As these studies are being pursued, we are in parallel producing the needed constructs for the ROSA26 locus. Should one or both loci prove to be Wnt responsive, and to faithfully recapitulate the known patterns of Wnt responsiveness in embryos, the system will have proven itself to be a valuable and simpler alternative to costly and time-intensive transgenic approaches.

The original system required mating the activator strain with the TRE-H2BGFP reporter strain; to eliminate the need for such breeding and subsequent genotyping for double transgenics, we now include the TRE-H2BGFP reporter with the SuperTOP-tTA2 activator, but oriented in the opposite direction to minimize potential expression problems. Cloning for the dual Activator-Reporter transgene is well underway. We are encouraged that this approach will work as we have observed that in cell lines, a single regulatable tTA2 is able to robustly activate a single TRE-H2BGFP reporter (4). We have also found that the level of tTA expression driven by the cytokeratin 14 (K14) promoter is sufficiently produced to activate the TRE-H2BGFP genes in our existing transgenic reporter line (Figure 6 and K14-tTA mouse line was provided by the Wysolmerski lab at Yale University). These studies provide precedent for the proposed approach.

Task 2: Determine whether mammary precursors exhibit functional characteristics of stem cells

While the last steps of the cloning and the targeting of the ROSA26 locus are being completed, we have made progress in establishing two key techniques required for this project. We successfully setup the mammary gland transplantation technique. As shown in figure 7A, a transplanted piece of adult mammary gland with endogenous expression of actin-EGFP successfully repopulated an epithelial-free #4 fat pad from an immunocompromised female mouse. Efficient repopulation of #4 fat pad was verified by both fluorescence detection of mammary epithelial ductal outgrowths (Figure 7) and whole mount carmine staining (not shown). We have also begun preliminary work on the transplantation of embryonic mammary buds. Repopulation of an adult mammary gland requires a minimum of 5 weeks before analysis, so we have learned an additional technique to expedite analysis: grafting embryonic mammary buds onto the renal capsule of immunocompromised female mice. Requiring only 3 weeks before verification of ductal outgrowth, this technique

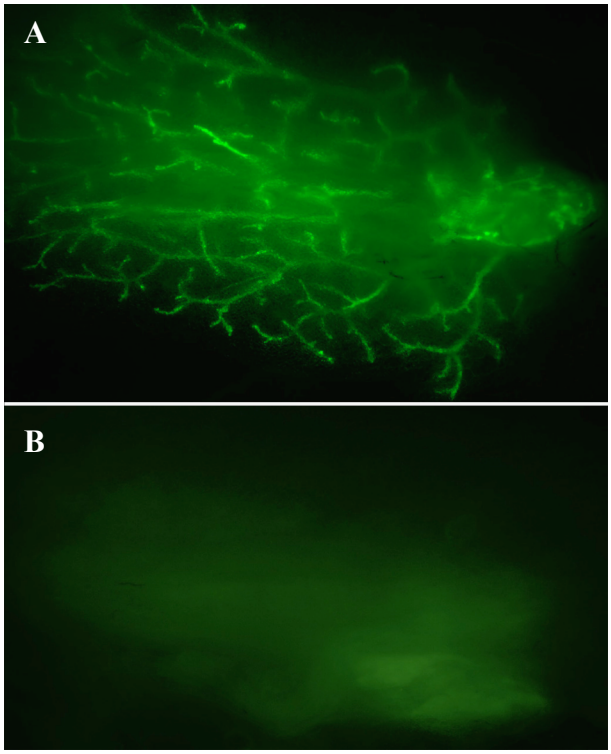


Figure 7A: Transplant of 1mm fragment of adult mammary gland #4 from actin-EGFP female into the right side, cleared fat pad #4 of a SCID female. **7B:** Sham – left side, cleared fat pad.

necessitates less time to determine whether our transplantation procedure was successful. After this procedure is optimized to maximize survival and efficiency, we will complete an analysis on dissociated embryonic mammary buds to master this technique before implementing this procedure for the evaluation of the stem cell characteristics of label retaining cells.

Task 3: Isolate embryonic mammary stem cells to define their gene expression profiles

An important long-term goal is to determine the gene expression patterns in embryonic mammary stem cells to determine which pathways are important for self-renewal and for maintaining their "stemness", and to then determine whether any of these pathways are evident in mammary cancers. We are first establishing the technique using embryonic mammary epithelial cells. This technique involves the dissection of a large pool of buds, dissociation into single cell suspensions, FACS to separate epithelial cells from the surrounding mesenchyme, isolation of RNA, and the amplification of RNA. Once we have optimized this procedure and verified that the amplification step does not substantially change the detected mRNA levels, we will work with our collaborators at the Salk in the laboratory of Dr. Ron Evans to do large scale quantitative PCR (qPCR) analysis on the nuclear receptor expression profile. We are entering into this collaboration since the Evans lab has pioneered the technique for quantitative analysis of nuclear hormone receptors. We are now developing primer sets to add to theirs to enable interrogation of known stem cell maintenance genes. We have chosen to utilize this qPCR profiling system rather than the more typical microarray analysis due to the higher reproducibility of the data, and its greater dynamic range, which enables quantitative differences in gene expression to be identified.
The initial studies will be performed on dissociated, unlabeled embryonic mammary buds. However, the genetic system described above will enable us to determine whether there exists in the early mammary buds, a subset of Wnt-responsive cells that exhibit characteristics of mammary stem cells. Upon Wnt activation of SuperTOP tTA2 and H2BGFP expression, H2BGFP will evenly decorate the chromosomes. If we then add doxycycline to prevent further H2BGFP expression, and a certain subset of the cells do not cycle frequently, they will remain more fluorescent than other cells in the population that do cycle. The bright cells would be the equivalent of label-retaining cells shown to correspond to stem cells in other systems. We will use FACS to isolate these bright cells, and determine if they have stem cell characteristics using mammary transplantation. If they do, then we will scale up the study, use FACS to isolate the bright cells after a doxycycline "chase", and then use the qPCR approach to determine their stem cell quantitative gene expression signature. The ultimate goal of such studies would be to identify molecular pathways that would be candidates for determining self-renewal, or limiting differentiation. The genetic systems we have proposed will then enable direct tests of which, if any, of the candidate genes are truly involved in stem cell maintenance, differentiation, and cancer.

KEY RESEARCH ACCOMPLISHMENTS

- Generation and analysis of five TOP-tTA Founders
- Proving the ability of the system to work in cells in and out of cycle
- Demonstrating ability of an analogous system to label early mammary buds
- Redesigning a modular system for site specific recombination into ES cells
- Targeting strategy for ROSA26 and ColA1 loci
- Establishing transplant procedure
- Establishment of multiple assays to detect Wnt responsiveness in individual cells and in whole embryos

REPORTABLE OUTCOMES

- T32 NIH Cancer Training Grant – post doctorate (joined 1/15/07)
- CMG NIH Pre-doctoral Training Grant – graduate student (joined 5/30/06)

CONCLUSION

We have validated a number of the underlying assumptions in the original proposal. However, we also identified problems that limited our ability to apply the initial experimental design. Our studies over the past year enabled us to develop a new and potentially far more simple and powerful labeling strategy for mammary stem cells. This more refined modular system avoids many of the conventional problems with traditional transgenesis. We are now positioned to test an essential component of the system, SuperTOP-tTA2-IRES-LacZ, in the ColA1 locus to evaluate Wnt activation and inducibility of this transgene. Once completed, the modular system will enable the labeling, isolation, and characterization of mouse mammary stem cells, and should facilitate elucidation of stem cells in other organs and within cancer models. This will permit the exploration of the relationship between tissue stem cells and cancer stem cells, allowing us to address many of the important questions within this rapidly expanding field that merges developmental and cancer biology.

REFERENCES

1. Al-Hajj M et al. Prospective identification of tumorigenic breast cancer cells. *PNAS*, 2003. 100 (7): 3983-3988.
2. Smalley, M and Ashworth, A. Stem cells and breast cancer: a field in transit. *Nature Reviews,* 2003. 3: 832-844.
3. Zambrowicz, BP et al. Disruption of overlapping transcripts in the ROSA beta geo 26 gene trap strain leads to widespread expression of beta-galactosidase in mouse embryos and hematopoietic cells. *PNAS*, 1997. 94 (8): 3789-94.
4. Wong, ET et al. Reproducible doxycycline-inducible transgene expression at specific loci generated by Cre-recombinase mediated cassette exchange. *NAR*, 2005. 33 (17): 1-13.
5. Beard, C et al. Efficient method to generate single-copy transgenic mice by site-specific integration in embryonic stem cells. *Genesis*, 2006. 44 (1): 23-8.
6. VP16 minimal domain vectors. Clontechniques, April 1998.
7. Baron, U et al. Tetracycline-controlled transcription in eukaryotes: novel transactivators with graded transactivation potential. *NAR*, 1997. 25 (14): 2723-2729.
8. DasGupta, R and Fuchs, E. Multiple roles for activated LEF/TCF transcription complexes during hair follicle development and differentiation. *Development*, 1999. 126: 4557-4568.